



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/660,780	09/12/2003	Nambi Seshadri	58268.00224	5880
32294	7590	04/27/2007	EXAMINER	
SQUIRE, SANDERS & DEMPSEY L.L.P.			LERNER, MARTIN	
14TH FLOOR			ART UNIT	PAPER NUMBER
8000 TOWERS CRESCENT				
TYSONS CORNER, VA 22182			2626	
SHORTENED STATUTORY PERIOD OF RESPONSE		MAIL DATE	DELIVERY MODE	
3 MONTHS		04/27/2007	PAPER	

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)
	10/660,780	SESHADRI, NAMBI
	Examiner	Art Unit
	Martin Lerner	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) Responsive to communication(s) filed on 06 February 2004.
 2a) This action is FINAL. 2b) This action is non-final.
 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) Claim(s) 1 to 15 is/are pending in the application.
 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
 5) Claim(s) _____ is/are allowed.
 6) Claim(s) 1 to 15 is/are rejected.
 7) Claim(s) _____ is/are objected to.
 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) The specification is objected to by the Examiner.
 10) The drawing(s) filed on 06 February 2004 is/are: a) accepted or b) objected to by the Examiner.
 Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
 Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) Notice of References Cited (PTO-892)
 2) Notice of Draftsperson's Patent Drawing Review (PTO-948)
 3) Information Disclosure Statement(s) (PTO/SB/08)
 Paper No(s)/Mail Date _____

4) Interview Summary (PTO-413)
 Paper No(s)/Mail Date. _____
 5) Notice of Informal Patent Application
 6) Other: _____

DETAILED ACTION

Drawings

1. The drawings are objected to because they are informal. Figures 1 to 5 contain handwritten elements and are hand drawn.

Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the examiner does not accept the changes, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Specification

2. The disclosure is objected to because of the following informalities:

On Page 10, ¶[0034], "detects" should be –detect—.

On Page 10, ¶[0036], "preliminary" should be –preliminarily—.

Appropriate correction is required.

Claim Rejections - 35 USC § 102

3. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

4. Claims 1 to 3, 5 to 7, and 9 to 11 rejected under 35 U.S.C. 102(e) as being anticipated by *Morris*.

Regarding independent claims 1, 5, and 9, *Morris* discloses a speech recognition method, device, and system, comprising:

“an audio signal receiver configured to receive audio signals from a speech source” – a user speaks to system 100, and system 100 captures the user’s speech with speech input unit 104 (column 4, lines 15 to 19: Figures 1 and 2: Block 202); speech is an audio signal;

“a video signal receiver configured to receive video signals from the speech source” – a user speaks to system 100, and system 100 captures the user’s image with video input unit 102 (column 4, lines 15 to 19: Figures 1 and 2: Block 202);

“a processing unit configured to process the audio signals and the video signals”

– system 100 combines any captured speech or video and proceeds to process the combined data stream in multi-sensor fusion/recognition unit 106 (column 4, lines 20 to 24: Figures 1 and 2: Block 204);

“a conversion unit configured to convert the audio signals and the video signals to recognizable information” – system 100 interprets any verbal input using the speech recognition functions of multi-sensor fusion/recognition unit 106; speech recognition is supplemented by visual information captured by video input unit 102, such as any interpreted facial expressions (e.g., lip-reading); a list of spoken words is generated from the verbal input (column 4, lines 25 to 31: Figures 1 and 2: Block 206); spoken words are recognizable information;

“an implementation unit configured to implement a task based on the recognizable information” – system 100 provides a response based upon whether the user has asked a question or made a statement; if a user has asked a question, then system 100 searches knowledge database 116 for a response to the objective question; a user may ask: “What is the weather in Phoenix, today?”; system 100 retrieves an answer, and the information is communicated as output via computer monitor and speakers (column 4, line 56 to column 5, line 24: Figure 3: Blocks 306, 308, 310, 312, 322); responding to a question by searching a knowledge database for a weather report for Phoenix, and outputting the weather report, is equivalent to implementing a task.

Regarding claims 2, 6, and 10, *Morris* discloses that video input unit 102 receives face/voice expressions and interpreted facial expressions including lip-reading (column 4, lines 27 to 30: Figures 1 and 2).

Regarding claims 3, 7, and 11, *Morris* discloses that, in one embodiment, processing by multi-sensor fusion recognition unit 106 is split into three parallel processes to minimize time of processing (column 4, lines 20 to 24: Figures 1 and 2).

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 13 to 15 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Morris* in view of *Houvener* ('588).

Concerning independent claims 13 to 15, *Morris* discloses a speech recognition method, device, and system, comprising:

"an audio signal receiver configured to receive audio signals from a speech source" – a user speaks to system 100, and system 100 captures the user's speech with speech input unit 104 (column 4, lines 15 to 19: Figures 1 and 2: Block 202); speech is an audio signal;

“a video signal receiver configured to receive video signals from the speech source” – a user speaks to system 100, and system 100 captures the user’s image with video input unit 102 (column 4, lines 15 to 19: Figures 1 and 2: Block 202);

“a first processing unit configured to process the audio signals” – system 100 combines any captured speech or video and proceeds to process the combined data stream in multi-sensor fusion/recognition unit 106 (column 4, lines 20 to 24: Figures 1 and 2: Block 204);

“a first conversion unit configured to convert the audio signals to recognizable information” – system 100 interprets any verbal input using the speech recognition functions of multi-sensor fusion/recognition unit 106; speech recognition is supplemented by visual information captured by video input unit 102, such as any interpreted facial expressions (e.g., lip-reading); a list of spoken words is generated from the verbal input (column 4, lines 25 to 31: Figures 1 and 2: Block 206); spoken words are recognizable information;

“an implementation unit configured to implement a task based on the recognizable information” – system 100 provides a response based upon whether the user has asked a question or made a statement; if a user has asked a question, then system 100 searches knowledge database 116 for a response to the objective question; a user may ask: “What is the weather in Phoenix, today?”, system 100 retrieves an answer, and the information is communicated as output via computer monitor and speakers (column 4, line 56 to column 5, line 24: Figure 3: Blocks 306, 308, 310, 312,

322); responding to a question by searching a knowledge database for a weather report for Phoenix, and outputting the weather report, is equivalent to implementing a task.

Concerning independent claims 13 to 15, *Morris* discloses “a second processing unit configured to process the video signal” and “a second conversion unit configured to convert the processed video signals into recognizable information” – system 100 interprets any verbal input using the speech recognition functions of multi-sensor fusion/recognition unit 106; speech recognition is supplemented by visual information captured by video input unit 102, such as any interpreted facial expressions (e.g., lip-reading); a list of spoken words is generated from the verbal input (column 4, lines 25 to 31: Figures 1 and 2: Block 206); spoken words are recognizable information. That is, multi-sensor fusion/recognition unit 106 performs the functions of processing the signals and converting the processed signals into recognizable information for the video signals as well as the audio signals. However, *Morris* omits a second processing unit that processes the video signals “when a segment of the audio signals can not be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information”. *Morris* has coinciding audio and video signals, but does not say what to do when an audio signal cannot be converted. Presumably, if the audio signals could not be converted into recognizable information, *Morris* would simply do the best it could with whatever information was present, including the video signals.

Concerning independent claims 13 to 15, however, *Houvener* ('588) teaches tiered biometric analysis, where a primary biometric data input unit receives primary

biometric data regarding a subject, and a secondary biometric data input unit receives secondary biometric data regarding the subject. If a primary biometric data match is below a minimum primary biometric data threshold, then the secondary biometric data input unit receives the secondary biometric data, and a secondary biometric analysis unit analyzes the secondary biometric data. (Page 1: ¶[0007]) Thus, *Houvener* ('588) suggests only utilizing secondary biometric data when the primary biometric data is below a threshold, corresponding to "when a segment of the signals cannot be converted into the recognizable information". Primary and secondary biometric data can include audio, voice, video, and images. (Page 2: ¶[0019] - ¶[0020]) An objective is to optimize the quality of the captured data presented to biometric analysis, and permit an operator to select the easiest to use biometric. (Page 5: ¶[0038]) It would have been obvious to one having ordinary skill in the art to employ the thresholding method for only utilizing secondary biometric signals when primary biometric signals are below a threshold and unrecognizable as taught by *Houvener* ('588) in a speech recognition method, device, and system for combining audio and video signals of *Morris* for a purpose of optimizing the quality of analogous art captured biometric data.

7. Claims 4, 8, and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Morris* in view of *Bakis et al.*

Morris does not expressly disclose a storage unit for storing the audio signals and the video signals to a destination source, and a transmitter for sending the audio signals and the video signals to a destination source. However, it is well known to

operate biometric identification via a client/server network, where biometric data is stored on a server, and biometric data is collected locally but compared to stored biometric data on the server. *Bakis et al.* teaches an analogous art method and apparatus for recognizing the identity of individuals by a speaker recognition system and a lip classifier, where biometric attributes are pre-stored for later retrieval so that they may be compared. Further, a server is included for interfacing with a plurality of biometric recognition systems to receive requests for biometric attributes therefrom and transmit biometric attributes thereto. The server has a memory device for storing the biometric attributes. (Column 8, Line 47 to Column 9, Line 16) Objectives are to provide a significant increase in the degree of accuracy of recognition and to provide a significant reduction in fraudulent or errant access to a service and/or facility. It would have been obvious to one having ordinary skill in the art to store and send biometric attributes to a server ("a destination source") as taught by *Bakis et al.* in a method, device, and system for combining audio and video signals of *Morris* for purposes of increasing accuracy of recognition and reducing fraudulent access.

Conclusion

8. The prior art made of record and not relied upon is considered pertinent to Applicant's disclosure.

Houvener ('536), Colmenarez et al., Chen et al., and Verma et al. disclose related art.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David R. Hudspeth can be reached on (571) 272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

ML
4/24/07



Martin Lerner
Examiner
Group Art Unit 2626